Interface Language, User Language and Success Rates in The European Library

Maria Gäde[1], Juliane Stiller[1], Richard Berendsen[2] and Vivien Petras[1]

[1]Berlin School of Library and Information Science, Humboldt-Universität zu Berlin
maria.gaede|juliane.stiller|vivien.petras@ibi.hu-berlin.de
[2]ILPS - ISLA, University of Amsterdam
richard.berendsen@gmail.com

Abstract. In this paper, TEL 2010 action logs are analyzed with a particular focus on the impact of language (user native language and interface language) on the success of a search session. Particular user actions are defined as success indicators for searches and sessions are divided into "successful" and "unsuccessful" sessions with respect to their outcomes. Two approaches for studying the impact of the language of the search interface are pursued: (1) the effect of concurrent language choice when associating the user language (determined by IP address) with the interface language and (2) the consequences of interface language changes during a session. The challenges of country and language identification via IP addresses are also discussed.

Keywords: LogCLEF, log file analysis, The European Library (TEL), interface language, success rate

1 Introduction

It seems to be generally acknowledged that the option to select the interface language according to the native or preferred language of the user is the first (and simplest step) in the process of adapting an information system for multilingual users. This system localization [5] has to be complemented by query and/or document translation features in order to be called a cross-language information system [10],[1]. Adapting the interface language to one's native language has consequences for the usability and familiarity of a system for a user, however, how much impact it has on the perceived success of a search session in a multilingual system is not determined.

This year's LogCLEF lab presents 3 tasks, one of which is to study the success of searches through log files. This paper studies the impact of the interface language on the perceived success of a search session using the TEL action logs for the year of 2010 pursuing 2 approaches: (1) the effect of concurrent language choice when the user sets the interface language to his or her native language (determined by user access country via IP address) and (2) the consequences of interface language changes during a session, assuming that the user changes the interface language with a goal towards improving the search experience.

The paper is organized as follows: section 2 summarizes related work on log file studies analyzing success rates of searches, while section 3 describes the TEL action log file. Section 4 looks at how language information and success indicators can be determined from log files and in particular the challenges of language identification from IP addresses. Section 5 discusses the findings analyzing sessions where interface language and user language coincide. Section 6 analyzes the findings comparing sessions where users change their interface language.

2 Related Work

Previous studies have investigated measures for success derived from log files. Huntington et al. [4] analyzed BBC search logs with respect to the number of searches conducted during sessions as well as lapse time between the searches of a session. They assume that users conduct many searches due to the fact that the retrieved results are not satisfying. A longer time period between searches for the same topic during a session is interpreted as extensive interaction with results and therefore as a satisfied information need. They found a relationship between searches that were limited to a scope and sessions with longer pauses.

Aula et al. [2] and Liu et al. [9] combined measures from log files with direct feedback from user studies. Aula et al. studied if and how the difficulty of search tasks influences search behavior. In contrast to other studies, the aim was to identify failure behavior patterns in order to support users facing complex search tasks. Users spend more time on result pages when facing difficult tasks and a more structured query refinement is representative for successful tasks. Liu et al. propose an examination of user behavior during query reformulation intervals (QRI) with regard to the usefulness of retrieved documents. The comparison of the duration of QRIs with and without saved documents showed that successful users spend more time interacting with search engine result pages (SERPs) and retrieved documents.

Kralisch [7] studied the impact of language and culture on user behavior in the electronic health domain. She investigated native and non-native speakers and their preference for different search options such as search engines, alphabetical search and hyperlink navigation. Through the analysis of log files she could not find a significant difference between native speakers and non-native speakers. Hassan [3] suggests a focus on successful or unsuccessful user goals rather than on document relevance.

Also TEL log files have been studied with respect to successful user paths or strategies. Lamm et al. [8] investigated user search performance and interaction for the The European Library (TEL) interface. They defined successful and not successful action patterns. A session without these actions and especially without a single full view is considered as failure.

Srinivasarao [11] analyzed action logs of 307 users and studied the paths of the most successful, the least successful as well the one between them according to the impact of language on search behavior. Success was characterized by the number of completed tasks (retrieved images). They found that the most successful users often reformulated their queries instead of looking at many result pages.

3 The TEL 2010 Action Logs

The TEL 2010 action log contains data about all user actions on the TEL portal (The European Library[1]) from January to December 2010. It logs user IP address, date & time, session information (i.e. an action belonging to a particular session), the selected interface language during the action, queries, other user actions (e.g. full view, save record, email record etc.) and result information (e.g. number of records retrieved, collections where records were retrieved from). In total, 25 actions are theoretically logged, but not all occur in the 2010 log file.
All lines (actions) from the log file were loaded into a MySQL database, where also all analyses were performed. The 2010 TEL log file had a total of 940,957 entries, which corresponded to 124,131 sessions. The following 15 actions were recorded (table 1):

Table 1. Action types for TEL logfile.

Action	Count	Description
view_brief	280618	short display of results
view_full	235171	detailed display of 1 result
search_sim	158339	search from simple search form
page_brief	126312	switch page in short result display
search_res	58094	search from results page
search_adv	36951	search from advanced search form
see_online	23956	see original object record
available_at	8097	link to original record (outside link)
service_all	4962	full record services link used
col_set_country	3303	search from collection browser
option_save_session_favorite	2435	save session favorite
jump_to_page	1788	jump in short result display pages
option_send_mail	718	email record
col_set_subj	152	search from subject collection browser
col_set_desc	32	description search for collection
none	29	no action recorded

In 93,185 sessions (out of 124,131), only the default interface language English is used. In 28,074 sessions, another interface language is used. All 36 offered interface languages were used[2] with Russian (3071 sessions), Portuguese (3022), French (2987), Polish (2451), German (2438), Turkish (1888), Spanish (1560), Greek (1322) and Italian (1270) as the most popular languages other than English.

[1] http://theeuropeanlibrary.org
[2]Logged are 38 different interface languages, but this must be due to renaming of language identifiers in the logfile.

There are 2872 sessions, where the interface language is changed by the user during the session. These were investigated further in section 6 as they contain reliable language information with respect to the user.

With the last 2 bytes of the IP addresses obscured, 11707 unique IP addresses were encountered in the 2010 log. Of those, 76% could be clearly associated with one country, the other 24% of IP addresses were ambiguous (more discussed in section 4.2).

4 Language Information and Success Rates from Logfiles

To investigate multilingual user behavior, several language aspects can be looked at: the native language of the user, the preferred retrieval language of the user, the interface language, the query language, document language etc. Gathering this language information from logfiles it not a trivial task, as most information that can be gleaned from logfile lines gives only implicit insight into language use or preferences. They do not necessarily determine the user's intention. There are indicators in the TEL logfile, which we consult to find answers to the following questions:

Table 2. Implicit and explicit indicators in the TEL action log.

Question	Indicator type	Action log
Native language of user	implicit	IP address
Search language of user	explicit	language of query
Preferred language for search results	implicit	language of results clicked language of query
Preferred portal / environment language	explicit	interface language change

For inferring the native language of users, the IP address and with it the country, where a user accesses the portal from, is used to determine this information. However, this indicator is only an approximation as users accessing from a particular country might not be native speakers of the national language or more than one language is spoken in this country (e.g. Switzerland).

However, even explicit language indicators like query language and interface language are difficult to determine. The language of a query, for example, can be ambiguous or non-determinable when named entities are searched [12]. The interface language cannot necessarily be determined from a http log and only an explicit interface language change indicated a language-conscious action of a user. It cannot be determined how many people prefer the default interface language and how many people just accept or "make do" with the default environment, because they either do not care to change the language or are not aware of this feature. Conversely, users also seem to change the interface language in the hopes of changing query or document languages, possibly a misunderstanding about the function of the interface language, which should be analyzed further.

4.1 Gathering Language Information from IP Addresses

Inferring user language information from IP addresses is at least a 2-step process: first, the IP address is converted to a geographic location or country (using IP address - country ranges published on the web) and second, determining the national or most spoken language in this country. For this paper, the IP-to-Country database by Webhosting was used to look up country information for IP addresses[3].

Due to privacy concerns, the IP addresses of all users in the logfile were partly obscured so that only the first 2 bytes were visible (e.g. 141.20.xxx.xxx). This caused some unforeseen challenges as not only the concrete location was obscured but even the country could not always unambiguously inferred from this information as the possible IP address range inferred from the first 2 bytes can span more than one country. One example would be the 2-byte range 121.58.xxx.xxx, where, if all possible IP addresses falling in this range were associated with a country, 144 would be identified with China, 64 with the Philippines, 16 with Japan, 16 with Australia, 8 with Indonesia and 8 with India. As different languages are spoken in those countries, a clear statement about the user language cannot be made.

Table 3 shows how many 2-byte IP addresses and consequently how many sessions can be associated with exactly 1 country: 76% and 75.6% respectively. This also means that of the 11707 2-byte IP address ranges obtained in the 2010 TEL logfile, ca. 24% could be associated with more than one country. For those 24%, which are ambiguous, one could as approximation select as most likely country (and therefore language) the one with the most frequent IP addresses falling in that range. In table 3, the largest range indicates the largest portion of an IP range assigned to a single country. In the session column, this indicates the certainty of a session being associated with this country.

Table 3. IP addresses to country association.

% largest range	User IP address (out of 11,707)[4]	Sessions (out of 124,131)
100%	76.0%	75.6%
[80% - 100%)	4.2%	2.3%
[60% - 80%)	2.1%	2.2%
[40% - 60%)	6.6%	7.8%
[20% - 40%)	7.6%	8.8%
[0% - 20%)	3.3%	3.5%

With these caveats in mind, user language information from IP addresses that are partly obscured, can only be rated as highly uncertain, both because of IP range to country conversion ambiguities and country to language conversion ambiguities. If

[3]http://ip-to-country.webhosting.info

[4]Note that 296 sessions have more than one user IP address - for those sessions we chose 1 IP address.

ambiguous IP address ranges are discarded, the analysis might be biased towards certain countries. It could be that ambiguous IP address ranges particularly skews the data with respect to small countries. To avoid this in future LogCLEF editions, a possibility may be to add a country code based on the full IP address to the log record before anonymizing the IP address.

4.2 Success Rate Indicators

In order to determine whether a session was successful for a user, indicators or actions that indicate a successful session need to be determined. Out of the 25 possible logged action types, we considered the following 5 as indicating when a user feels he or she might have reached a goal or fulfilled an information need:

- view_full (1 record is looked at in detail)
- see_online (original object at native interface is looked at, only visible when thumbnail is provided)
- available_at (link clicked to display metadata or object at original library site)
- option_save_session_favorite (record is saved by user)
- option_send_mail (record is emailed by user).

For a stricter division, we can also divide the first 3 actions as softer indicators (users might realize they do not need this record after all) and the last 2 actions as harder indicators (the user definitely considers this record relevant). A session is successful, if one of these indicator actions occurs. These actions nevertheless have to be reviewed in context. For the hard indicators it is obvious that a soft indicator action would have to precede in a session. The action logs do not always show this sequence and hard indicators may occur without the actions that should naturally occur before them. Additionally, the choices for actions presented to the user vary depending on provider and object looked at. Not all of them present the option of the ‘see_online’ button and the ‘available_at’ link so that some sessions may not show these indicator actions because they never appeared to the user.

5 User Language and Interface Language

We first show how the sessions for which we know the country of access for the user (75.6% of the sessions with a 100% certain IP address conversion rate) are distributed over countries

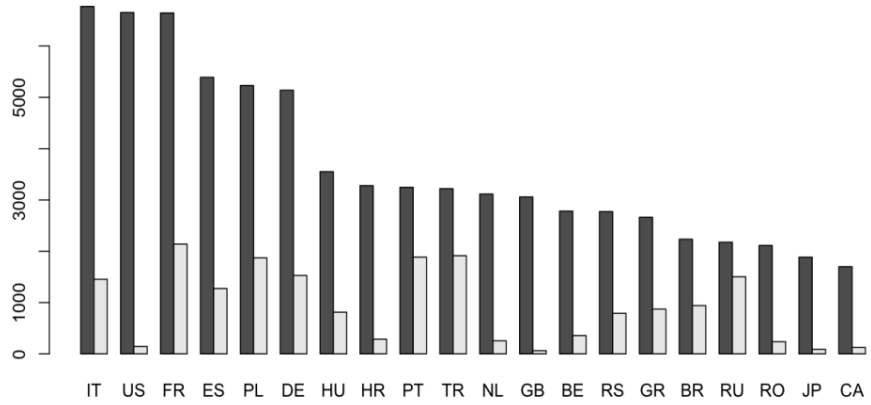

Fig. 1. Sessions per country.

We can also order these top 20 countries by the percentage of sessions in which an action with a non English interface language occurred (figure 2). The dark bar represents the number of sessions overall accessed from this country, the grey bar of sessions that are accessed (at least partly) with an interface language other than English. Users in Russia seem to use TEL most often with a non default interface language, closely followed by Turkey and Portugal. As this analysis contains only 75.6% of the sessions that could be unambiguously mapped to one country, other countries might not be considered in this distribution.

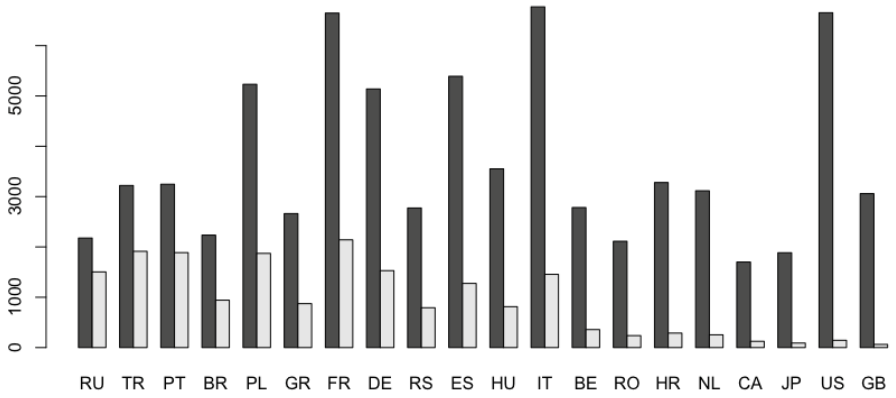

Fig. 2. Sessions with only English or mixed interface language per country.

The sessions are now split into two groups: one group that contains at least one action with a non English interface, and one group that contains only actions with the default interface language. We compare these sessions with respect to their success rates.

Table 4. Occurrence of hard and soft indicators per session.

Success action	**At least 1 non English 30945 (25%)**	**Only English 93185 (75%)**	**All sessions 124,131 (100%)**
Soft indicators	13808 (44.6%)	40518 (43.5%)	54326 (43.8%)
view_full	13339 (43.1%)	38874 (41.7%)	52213 (42.1%)
see_online	2578 (8.3%)	7098 (7.6%)	9676 (7.8%)
available_at	1765 (5.7%)	3291 (3.5%)	5056 (4.1%)
Hard indicators	723 (2.3%)	1207 (1.3%)	1930 (1.6%)
option_save_session_favorite	651 (2.1%)	978 (1.0%)	1629 (1.3%)
option_send_mail	117 (0.3%)	297 (0.3%)	414 (0.3%)
All indicators	13819 (44.7%)	40544 (43.5%)	54363 (43.8%)

In table 4, one cannot observe a big difference between sessions where an interface language other than English occurs and sessions with an English interface only, however, sessions that contain at least one non English action are a bit more successful in terms of the hard indicators. There are many possible causes for this that we could investigate in future work. For example, it could be that these sessions are typically longer and users engage more with the system.

If English is the only interface language used in a session, this could mean the people prefer or accept to experience their environment in English, because it's their native language, because they feel comfortable with it as a second language, because they do not want to change the interface language or because they do not know how to change it. A large part of this group may speak English well enough as a second language, but surely not everyone. If we could isolate the non English speaking user group, we can answer the question: 'Does it help if a search engine speaks my language?'. Therefore we now turn our attention to the 75.6% sessions for which we know the country.

We first establish a list of countries where the primary language is English. We want to exclude these countries from our analysis. According to the Wikipedia page on the English language, it is the official language in 51 countries, the de facto language for six more countries, and then a language spoken in a long list of non sovereign entities, most small, from which we selected the two largest: Hong Kong and Puerto Rico. After discarding all sessions from these 59 countries we have a set of 80565 sessions left to analyze. Now we can compare those sessions of seemingly non English native speakers who change the interface language and who leave or set English as interface language.

Table 5. Occurrence of hard and soft indicators per session from non English native speakers.

Success action	**Non English 21775 (27%)**	**Only English 58790 (73%)**	**All sessions 80565**
Soft indicators	9807 (45.0%)	26059 (44.3%)	35866 (44.5%)
view_full	9471 (43.5%)	25005 (42.5%)	34476 (42.8%)
see_online	1837 (8.4%)	4539 (7.7%)	6376 (7.9%)
available_at	1055 (4.8%)	2190 (3.7%)	3245 (4.0%)
Hard indicators	479 (2.2%)	722 (1.2%)	1201 (1.5%)
option_save_session_favorite	430 (2.0%)	585 (1.0%)	1015 (1.3%)
option_send_mail	80 (0.4%)	189 (0.3%)	269 (0.3%)
All indicators	9812 (45.0%)	26067 (44.3%)	35879 (44.5%)

In table 5, we see the same trend as above with seemingly no overall differences, although there are more non English interface language sessions that have a hard indicator action for success. Since also this table disregards all sessions for which a country of access could not be unambiguously identified (ca. 25%), a bias might have been introduced.

6 Interface Language Changes and Success Rates

Possibly the only reliable information which can be used to draw conclusion about the users' preferred language is the interface language switch. This section focuses on sessions where users changed the interface language and are therefore different from sessions with a default English language interface.

The European Library allows the user to select the interface language by a drop down menu. In the action logs, this switch is logged and the language of the interface is provided for any action logged. The deployment of stateless URLs (REST architecture) forwards even a returning user (who might have changed preferences before) to the default English interface. A direct non-English entry is only possible by bookmarking the appropriate language version or through an search engine result page (SERP) link.

Consequently, three session types with respect to interface language can be differentiated:

- no interface language change (default language English)
- direct access with non-English interface language (bookmark or SERP)
- language change during session (via interface).

6.1 Interface Language Switches

Most users do not perform a language change (75% of all sessions in the TEL 2010 action log), some choose direct access (23%) and only very few (2% or 2872)

consciously change the interface language during their session. When a change in interface language occurs, two types can be observed:

- one interface language change during the session (2700 sessions or 94%),
- several interface language changes during the session (172 or 6%): out of those, most users change the interface language twice (122) or three times (31).

A particular sub type of language switch occurs, when the interface language is changed several times during the session, however, the user switches back and forth between two languages, e.g.: en → de → en. As a special case, this occurs 142 times or about 5% of the time.

Overwhelmingly, users switch from the default English interface language to their preferred language (75%), however, also other languages are changed (when a different language version was bookmarked or searched for and the user switches to another language).

6.2 Success Rates in Interface Language Change Sessions

The 2872 sessions, which contain an interface language change, were further analyzed by actions occurring before and after the language change. For this, the sessions were split up into smaller sessions each containing entries with one specific interface language.

Almost half of the sessions (1351 out of 2872) contained only one action before the interface language was changed. Most users (80%) change immediately after conducting a simple search and continue again with a simple search in their preferred language.

Comparing the actions conducted before and after the first interface language change (table 6), one can observe that the frequency of any particular action increases after the language change, however, the frequency distribution of actions does not change (with 1 exception: option_save_session_favorite, where both the frequency and therefore the order in the frequency distribution changes).

Table 6. Actions performed before and after language switch.

Action	Before language switch	After first language switch[5]
view_brief	4743	6770
view_full	4291	5552
page_brief	2404	3203
search_sim	4350	3083
search_res	1134	1990
see_online	776	908
search_adv	705	781
available_at	189	241
service_all	189	222
col_set_country	143	140
option_save_session_favorite	84	70
jump_to_page	58	87
option_send_mail	9	26
col_set_subj	8	12

The following table (7) compares success indicators within interface language change sessions, before and after the first language switch and all sessions.

Table 7. Comparison success indicators within all sessions and sessions with LC.

Success actions*	Sessions before LC (2872)	Sessions after first LC[6] (2872)	Sessions with LC (2872)	All sessions (124,131)
Soft indicators	724 (25,2%)	1024 (35,7%)	1674 (58,3%)	54326 (43,8%)
view_full	648 (22,6%)	902 (31,4%)	1622 (56,5%)	52213 (42,1%)
see_online	209 (7,3%)	289 (10,1%)	528 (18,4%)	9676 (7,8%)
available_at	84 (2,9%)	151 (5,3%)	241 (8,4%)	5056 (4,1%)
Hard indicators	68 (2,4%)	71 (2,5%)	138 (4,8%)	1930 (1,6%)
option_save_session_favorite	59 (2%)	55 (1,9%)	121 (4,2%)	1629 (1,3%)
option_send_mail	9 (0,3%)	16 (0,6%)	26 (0,9%)	414 (0,3%)
All indicators	739 (25,7%)	1034 (36%)	1674 (58,3%)	54363 (43,8%)

Comparing sessions with an interface language change to the whole action log corpus, sessions with a language change do seem to be slightly more successful. One possible explanation for this outcome is that if a user switches the interface language,

5 Actions after the second language change are not considered.

6 Actions after the second language change are not considered.

* Sessions with at least one success indicator.

he or she is engaging with the website and more likely to spent more time on it and therefore probably more successful.

The whole corpus has 39090 sessions, which consist only of a single action and are therefore more unlikely to be successful. Sessions with a language change contain 16 actions on average whereas all sessions contain only 8 actions on average.

7 Conclusion

In this paper, we presented two different approaches to study the impact of language on the user's success within a session. We derived language information from the IP address and the interface language change. Both indicators are rather weak as our analysis shows.

IP addresses obscured to two bytes created an uncertainty for 25% of the unique IPs we tested. Only 75% of the action log sessions for this year's LogCLEF corpus could be assigned to their respective country. Further analysis needs to be done in order to determine which countries are easy to assign, which are most likely ambiguous and whether their is a bias making small countries harder to detect.

The second approach analyzed the impact of interface language change on success indicators finding an increased number of success actions for these sessions compared to the complete corpus. Since users rarely switch the interface language (2%), this analysis only includes a small subset of sessions and users. Nevertheless, the interface language switch seemed to be the most reliable user input regarding their preferred language.

During our analysis we encountered many problems with respect to analyzing the user's preferred language. The information which can be drawn from TEL action logs are limited. The only reliable input about the users' preferences is the language interface change. The user has to actively engage with the system and choose the desired language from a drop-down menu. Although this seems to be a straightforward expression of the user's choice it is not clear what she or he intended with it. On average, the language is only switched after 7 actions. This could mean that the user either did not find the drop-down menu before or was unsatisfied with the results and thought this might influence the result set. Whatever the interpretation, it shows that any conclusions drawn from the language interface change might be incorrect. Future work could also include other language information such as the language of the user agent or language of the results clicked. This would require that this information is logged as well.

We encountered many problems which were due to flaws in the data. It seems that for the action logs it is not possible to properly reconstruct user entry points. As stated above, over 30% of the sessions have only one action, some of them being the action 'full_view', which is an impossible entry point for the portal. By using Google Analytics[7] cookies for reconstructing sessions, the user path and clickstream might be restored more reliably making an analysis of user language preferences more meaningful.

[7]http://analytics.google.com

Acknowledgement. We would like to thank Sjoerd Siebinga for his support with the IP address to country conversion tables.

References

1. Agosti, M.; Crivellari, F.; Deambrosis, G.; Ferro, N.; Gäde, M.; Petras, V. and Stiller, J. (2009). D2.1.1 – Report on User Preferences and Information Retrieval Scenarios for Multilingual Access in Europeana. EuropeanaConnect ECP-2008-DILI-528001. http://www.europeanaconnect.eu/documents/D2.1.1_eConnect_Report_User_Preferences_MLIA_v1.0_20091222.zip
2. Aula, A.; Khan, R.M. and Guan, Z. (2010). How does search behavior change as search becomes more difficult?. In: Proceedings of the 28th international conference on Human factors in computing systems (CHI '10). ACM, New York, NY, USA, pp. 35-44.
3. Hassan, A., Jones, R. and Klinkner, K. L. (2010). Beyond DCG: User behavior as a predictor of a successful search. In: WSDM '10: Proceedings of the third ACM international conference on web search and data mining. ACM.
4. Huntington, P.; Nicholas, D.; Jamali, H.R. (2007). Employing log metrics to evaluate search behaviour and success: case study BBC search engine. In: Journal of Information Science 33(5): 584-597.
5. Ishida, R. and Miller, S. K. (2006). Localization vs. Internationalization.
6. http://www.w3.org/International/questions/qa-i18n.
7. Kralisch, A. (2006). The Impact of Culture and Language on the Use of the Internet: Empirical Analyses of Behaviour and Attitudes. Doctoral Thesis, Humboldt-Universität zu Berlin, Germany. http://edoc.hu-berlin.de/docviews/abstract.php?id=27410
8. Lamm, K.; Mandl, T. and Koelle, R. (2009). Search path visualization and session performance evaluation with log files. In: Peters, C.; Di Nunzio, G.; Kurimo, M.; Mandl, T. and Mostefa, D. (Eds.). CLEF 2009, Springer-Verlag: Berlin, Heidelberg, pp. 538-543.
9. Liu, C.; Gwizdka, J. and Liu, J. (2010). Helping Identify When Users Find Useful Documents:Examination of Query Reformulation Intervals. In: Proceedings of IIiX 2010, August 18-21, 2010, New Brunswick, New Jersey, USA, pp. 215-224.
10. Oard, D. W. (2006). Transcending the Tower of Babel: Supporting Access to Multilingual Information with Cross-Language Information Retrieval. In: Popp, R. and Yen, J. (Eds.) Emergent Information Technologies and Enabling Policies for Counter-Terrorism. Prentice Hall: Upper Saddle River (NJ), USA, pp. 131–139.
11. Srinivasarao, V. (2008). Mining the Behavior of Users in a Multilingual Information Access Task. In: CLEF 2008 Working Notes. http://www.clef-campaign.org/2008/working_notes/vundavalli-paperCLEF2008.pdf
12. Stiller, J., Gäde, M., Petras, V. (2010). Ambiguity of Queries and the Challenges for Query Language Detection. CLEF 2010 LogCLEF Workshop. In: CLEF 2010 Labs and Workshops Notebook Papers. M. Braschler, D. Harman and E. Pianta. Padua, Italy, 22-23 September 2010.